

Improvement Strategies for State Performance Planning SPP Toolkit Series

Number 1

A Few Steps to Better Data

Matthew Klare
National Dropout Prevention Center for
Students with Disabilities, Clemson University



National Dropout Prevention Center
for Students with Disabilities



U.S. Office of Special
Education Programs

November 2006. This information is copyright free. Readers are encouraged to copy and share it, but please credit the National Dropout Prevention Center for Students with Disabilities (NDPC-SD) www.ndpc-sd.org. Publication of this document is made possible through a Cooperative Agreement between the National Dropout Prevention Center for Students with Disabilities and the Office of Special Education Programs of the U.S. Department of Education. The contents of this document do not necessarily reflect the views or policies of the Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. Dr. Selete Avoke is the OSEP Project Officer.

For further information about NDPC-SD, please contact:
Loujeania Williams Bost, Ph.D., Director
National Dropout Prevention Center for Students with Disabilities
College of Health, Education, and Human Development
Clemson University
209 Martin Street
Clemson, SC 29631-1555
Telephone: (800) 443-6392
(864) 656-2599
TDD/TDY: (866) 212-2775
Fax: (864) 656-0136
www.ndpc-sd.org

NDPC-SD was established to support state education agencies in assisting local education agencies to increase school completion rates and decrease dropout rates among students with disabilities. NDPC-SD supports states through the following activities:

- Identifying evidence-based dropout prevention interventions, programs and practices.
- Producing evidence-based knowledge useful to school professionals.
- Providing targeted technical assistance to states in a variety of formats.
- Disseminating dropout prevention information through multiple methods.

Acknowledgements

NDPC-SD would like to thank Marsha Braun of Westat, Jane Sullivan of the Pennsylvania Department of Education, José Blackorby of SRI, Inc., and Loujeania Bost of NDPC-SD for their thoughtful reviews and insightful suggestions for this guide. We would additionally like to thank our OSEP project officer, Dr. Selete Avoke, for his continued guidance and support of the Center's work.

Introduction

In order to comply with the Individuals with Disabilities Education Act (IDEA) 2004, states must collect valid data as needed to report annually to the Secretary of Education. This requirement, combined with those set forth in the No Child Left Behind (NCLB) legislation, has led states to collect and analyze data about their programs, districts, schools, and student performance and take actions based upon the results. The proliferation of educational data and the increased dependency on those data have put educators in a challenging position. Because critical decisions are made and actions taken based on the data, it is imperative that these data are reliable, valid, and correctly interpreted.

As evidenced in the State Performance Plans and discussed in the National Governor's Association's *Graduation Counts* document, high school graduation and dropout data are often not collected in a consistent manner, nor are the data reliable. Until recently, many states did not collect both graduation and dropout data. Of those that did, few stored the data in longitudinal databases. For these reasons, many states have been unable to track all of their students as they progressed through high school.

To improve the state of affairs of high school data, the procedures used in collecting, analyzing, and comparing the data should be consistent. This will permit educators to make valid comparisons as well as assessments of states' performance over time. Insights gained through such comparisons can help educators and stakeholders identify needs as well as strengths that can be capitalized upon to make positive changes that improve outcomes for youth with Individualized Education Programs (IEPs).

The need for valid data to guide educational improvements for youth with IEPs cannot be overstated. However, in developing Annual Performance Reports (APRs) and State Performance Plans (SPPs), many states reported difficulties when working with their dropout and graduation data—two of the Office of Special Education (OSEP) Program's monitoring priority indicators. Some of the difficulties arose in calculating and comparing dropout and graduation rates for all students and youth with IEPs.



The procedures used in collecting, analyzing, and comparing the data should be consistent in order to support valid comparisons as well as assessments of states' performance over time.



While the problems first manifested themselves at the state level, they were magnified at the national level when comparisons were made across the 60 states, territories, and agencies, many of which employed different definitions and methods.

Some of the issues described in this document are within the control of states. These include the type of dropout or graduation calculations employed by each state, the exact dates of data collections, and the manner in which educational data are stored. Other factors have arisen from differences in definitions and requirements specific to various federal laws and regulations. The purpose of this practice guide is to discuss several of the factors that caused difficulties with states' data and suggest some practical solutions. The content is designed for state education agency personnel, including special education directors, data managers, and program developers.

The Malady: Factors That Cause the States Difficulty

Seven major factors impact a state's ability to make effective use of its indicator data for dropout and graduation. These factors are:

- the quality of district-level data, especially the granularity of the data collected for general-education and youth with IEPs;
- the comparisons made when calculating dropout and graduation rates;
- the timing of general enrollment and student-level data collection;

- the definitions of “dropout” and “graduate”;
- the actual calculations used in arriving at these rates;
- the manner in which states store educational data; and
- the underutilization of the data.

1. **Data quality.** The quality of the data received from districts is a foundational issue. Student-level and general enrollment data submitted by districts can vary in quality. For instance, some districts consistently follow proper procedures, perform necessary checks, and submit their data on time, whereas other districts do not. Some of this variability is a function of the training and support the state provides to districts, while some is due to funding and personnel issues in districts. Regardless of its cause, if local and state agencies start with bad data, everything that relies on the data will be flawed.
2. **Comparison errors.** The second major factor at the state level is the type of comparisons made when states calculate dropout and graduation rates. Some of the variability observed across states can be accounted for through their use of the wrong student data while making their comparisons. For example, states were asked to compare data for youth with IEPs to all student data when calculating dropout and graduation rates for the SPP. While many states made this comparison, some used counts of nondisabled students in the calculations instead of all student data (which includes youth with IEPs). The result was an underrepresentation of the actual dropout and graduation rates for all these states.
3. **Timing of data collections.** The timing of data collections also contributes to discrepancies between graduation rates and dropout rates for all students versus youth with IEPs. For example, data for youth with IEPs are generally derived from the 618 Data Collection, which has traditionally occurred on December 1 of the year, whereas all student enrollment data were generally collected earlier in the fall, generally in September or October. This difference in timing has reduced the comparability of the data, thereby decreasing the validity of comparisons made between youth


with IEPs and their nondisabled peers. OSEP requires states to collect exiting data for the period from July to June. States are now permitted to collect their child count and Least Restrictive Environment (LRE) data on any date from October 1 through December 1, which has the potential to help alleviate the problem.

4. **Definitions of terms.** Some of the difficulties associated with quantifying dropout and graduation rates can be attributed to a lack of standardized definitions of what constitutes a graduate or a dropout. Several factors prevent a clear definition. Among these are the variability in the age group or grade level of students included in the calculations and the inclusion or exclusion of particular groups or classes of students (e.g., students receiving a GED or occupational diploma) from consideration in the calculations.

For example, NCLB defines the graduation rate as the percentage of students who graduate from secondary school with a regular diploma in the standard number of years. This definition is employed in calculating all student graduation rates, yet under federal law, youth with IEPs are eligible to be served by schools until age 21, and some states also extend eligibility for services beyond 21, so they frequently fall outside of the strict NCLB definition of “standard number of years,” which excludes them from the calculation.

Another point of confusion is the definition of what constitutes graduation with a “regular diploma.” Numerous states offer multiple diplomas. In some states, recipients of these diplomas are counted as “regular” graduates, whereas, in other states they are not. Some states have occupational diplomas that are considered “regular” although the associated curricula involve no regular education classes. Additionally, some states include students who earn a General Education Development (GED) degree in their graduation rate calculation, while other states count them as dropouts.

General education enrollment is usually described by grade levels, whereas enrollment information for youth with IEPs is usually reported as a function of student age. This makes sense, but it confounds



clear comparisons between the two groups of students. An example of this occurred when states were developing the SPPs in the fall of 2005. In order to achieve a semblance of comparability between the dropout and graduation rates for youth with IEPs and their general education peers, many states used the 618 data and assigned youth with IEPs to grades based upon their age. While it allowed for comparisons, it was not a particularly safe assumption for this group of students.

Another age-related issue is that many states include students from ages 14–21 in their dropout calculations, whereas other states include only those students who are ages 17–21. Some of the other definitional differences are more subtle than this, but in the end, the result of these differences serves to diminish meaningful comparisons of dropout and graduation rates across states.

Additionally, the codes in the data system that indicate a student's reason for no longer being enrolled (leaver codes) cause an inordinate amount of confusion and discrepancy in the calculation of dropout rates. For example, the “moved not known to be continuing” category is a major source of error in dropout calculations. The IDEA exiting data show that 78,146 students between the ages of 14 and 21 dropped out during the 2003-04 school year. During that same period, 43,914 students within the same age bracket fell into the “moved not known to be continuing” category of exiting—i.e., 43,914 students could not be accounted for. While their numbers are reported separately from dropout in the 618 dataset, OSEP includes these students as dropouts. If states were able to account for these students, the dropout rate would be reduced, but more importantly, it would reflect a more accurate picture of the number of youth with IEPs who drop out.

5. **Methods of calculation.** Another factor hampering valid comparisons is differences in the method of calculating dropout and graduation rates. Three methods are commonly used to calculate dropout rates; two are commonly used in calculating graduation rates. Different states employ the various methods, though slight variations exist in these methods from state to state. Additionally,



The “moved not known to be continuing” category is a major source of error in dropout calculations. The IDEA exiting data show that 78,146 students between the ages of 14 and 21 dropped out during the 2003-04 school year. During that same period, 43,914 students within the same age bracket fell into the “moved not known to be continuing” category.



different calculations are sometimes used to determine a state's youth with IEPs and its all student dropout and graduation rates.

The three basic types of dropout rates are: event rates, status rates (“leaver” rates), and cohort rates. Event rates indicate the percentage of students who drop out in a single year. Status rates reflect the percentage of students of a given age who have not finished high school or are not enrolled at a particular point in time. Cohort rates indicate the percentage of a single cohort of students (e.g., the students who entered ninth grade in the fall of 2003) who drop out over time (usually the four-year period associated with a “standard” high school education).

In general, states employing an event calculation tend to arrive at lower dropout rates than states using a leaver or cohort-based calculation. Event rates are calculated by dividing the number of students in a particular category or grade who drop out during a single year by the overall number of students in that category or grade who are enrolled. Status and cohort rates reflect multiple years of

attrition from high school, rather than the single year described by an event rate, so these tend to be higher than event rates. Because the cohort rate has a longitudinal component in which an actual group of students is followed through high school, it provides a more accurate picture of the nature of attrition from high school over the course of four years than do the other methods.

The two flavors of graduation rate calculations are the OSEP method and that recommended by National Center on Educational Statistics (NCES). The OSEP formula provides a single-year rate of graduation for a particular year, whereas the NCES formula reflects a graduation rate for a four-year cohort of students.

Graduation rates calculated using the OSEP formula cannot properly be compared with those derived using the NCES formula. The OSEP formula provides a snapshot of the graduation rate for a particular year, disregarding attrition. This tends to yield a higher graduation rate than does the cohort method, which yields the percentage of students who completed four years of high school and graduated.

6. **Data collection and storage.** The collection and storage of student data play a large role in determining what comparisons can validly be made, as well as the ease with which they can be made. In addition, the granularity of the data—the level of detail represented in the data—also affects the types of analyses and comparisons that can be performed. Typically, data for youth with IEPs (618) are more detailed than are general enrollment data. This difference in granularity prevents states from being able to disaggregate some of their general education data.

For example, many states are currently not able to calculate cohort dropout or graduation rates because they are not collecting enrollment data at the individual student level. This is a case of the data not being sufficiently detailed or granular to support the desired analyses. If data hasn't been collected, it can't be analyzed!

Data for the general student population and youth with IEPs have generally been stored in


different databases that employ different data fields and layouts. So, in addition to incompatibilities created by collecting data at different times of the year, data records for these students were further separated from each other by differences in their milieu of storage as well as structural differences in the records themselves.

Traditionally, many states' general enrollment databases have not supported data records on an individual student basis, making it impossible to keep track of a student if he or she moved from one school district in the state to another. On the other hand, data collected for youth with IEPs have generally been collected and stored at the individual student level. Even for youth with IEPs, most data systems lacked the ability to follow individual students easily if they moved between school districts.

These structural and philosophical differences in the approach to data collection made it impossible for many states to calculate true cohort rates of dropout or graduation for the general student population. At a national level, these differences have confounded our ability to make accurate comparisons of dropout and graduation rates for youth with IEPs and general education.

Recently, the U.S. Department of Education funded several states in the development of longitudinal databases that support data at an individual student level and provide connections between student demographic information, test scores, and other achievement-related data. The development of such databases will be a major step in the convergence of data for youth with IEPs and general education data. If we hope to be able to make accurate comparisons of information about these two groups of students, the development and careful use of a common set of data is essential.

7. **Underutilization of data.** While educators are collecting and analyzing more data than ever before, they are not necessarily maximizing the return on their investment of time, effort, and dollars. The final report of the Study of State and Local Implementation and Impact of the Individuals with Disabilities Education Act (SLIIDEA)—a study commissioned by OSEP to



determine the extent to which states, districts, and schools have implemented the 1997 Amendments to IDEA—dramatically illustrates the extent to which the educational community leverages existing data. According to this report, when data were available to districts and schools, they were more likely to be employed in program evaluation than in planning professional development or other activities. Eighty-three percent of districts and 88% of schools used data on academic performance for program evaluation. Considerably fewer districts (53%) and schools (49%) used the data for planning professional development.

Another example of the problem was districts' and schools' use of dropout and graduation data. The study showed that 85% of districts and 86% of schools employed graduation or dropout data for the purposes of program evaluation. However, only 30% of districts and schools applied this information for the purposes of professional development. Additionally, 42% of districts and schools incorporated data about parental involvement in the program; however, only 16% of the districts and schools factored what the data showed into professional development planning or implementation.

The Prescription—Some Practical Solutions

Key in addressing these issues is the need for consistency both in definitions of terms and in practices employed in the collection, analysis, and reporting of data. Ensuring that all parties involved in the data chain operate using the same set of terms and procedures will reduce extraneous errors and contribute to the fidelity of states' data.


Ensure internal consistency. States need to standardize the processes involved in collecting, analyzing, reporting, and using dropout and graduation data. These efforts should be focused both at the local/district level as well as at the state level. Choose data tools and procedures that will give the most accurate results.

Make the same comparisons for youth with IEPs and their nondisabled peers. To ensure that comparisons are valid and meaningful, make the same comparisons


for youth with IEPs that are made for general education students. For example, the enrollment data being compared should reflect as close to the same time period as possible for all students as for the youth with IEPs. Minimizing the differences in definitions and methods between the two groups will increase the accuracy of the comparisons that are made.

Follow a consistent set of procedures and data definitions. As a starting point in addressing this situation, all districts within a state should follow a consistent set of procedures in organizing and maintaining their enrollment databases. The use of a consistent data layout, as well as common field names and attributes, is also fundamental. For example, to the extent possible, the districts in a state should employ a standardized set of course numbers, course names, and course descriptions. This would make it easier to determine eligibility for graduation as well as facilitate placement and other decisions when students move from one district to another.

Ensure compatibility of databases. Data for general education and youth with IEPs should be maintained in compatible databases. Because the data collected about youth with IEPs are more detailed than that for a general education student, it may not be practical to store data for both groups in a single database. Recognizing this fact, it is still necessary to structure the two databases so they share as many fields and characteristics as possible. This will foster interoperability and help support comparisons of the two groups of students once the data have been collected.



Ensuring that all parties involved in the data chain operate using the same set of terms and procedures will reduce extraneous errors and contribute to the fidelity of states' data.



Self-Check: Can Your State Answer “Yes” to These Questions?

✓	Are the same calculations used to produce the all student rates and youth with IEPs rates of dropout and of graduation?
✓	If your state is not already using a cohort method for calculating dropout and graduation rates, is it moving in that direction?
✓	Are data for all students and youth with IEPs comparable? Definitions, calculations, and methods used should be compatible.
✓	Can the state’s database track individual students as they move through high school (even if they move to another school or district)?
✓	Does the state analyze data for one or more indicators and apply what is learned to reducing the dropout rate and increasing the graduation rate?
✓	Is the state using the data for planning and decision making?

Maximize Your Efficacy

Designing procedures and systems to support current and future state needs will allow states to calculate more accurate dropout and school completion rates. An integral part of the design process is careful attention to detail, not only of data elements that need to be collected, but of data security and privacy issues.

Use cohort rates to describe graduation and dropout.

Cohort rates optimally show how many students actually complete high school and how many don’t make it. Their use in describing graduation and dropout will yield more accurate and meaningful school completion and dropout rates than will the use of other calculations. While it takes four years before data for the first cohort become available, the quality of data will offset the wait.

The major prerequisite to using a cohort method is having a unique, statewide student identification system. Having a unique identifier for each student in the state makes it possible to follow each student’s progress through school both from year to year and across district boundaries.

While the use of such identification numbers has raised privacy issues, the practice is essential in order to facilitate the calculation of true dropout and

graduation rates. If basic precautions are employed in developing unique student identifiers, maintaining students’ privacy should be an attainable goal.

Choose or design a data system that meets current and anticipated future needs of the state. Before buying or building a data system, you should thoroughly understand your data needs and objectives. In developing this understanding, consider the following questions.

- Who will use the data and for what purpose(s)?
- How and by whom will data be entered into the system, accessed, or distributed?
- Does the system support predefined entry forms and queries that make it user friendly—even for relatively unsophisticated users?
- What level of granularity will be needed for the various types of data to be collected?
- In what format or formats will the data be collected, stored, compared, or analyzed?
- What analyses, comparisons, disaggregations, and reports will be needed?
- What policies and procedures must be instituted to ensure data consistency and quality?

- What might the state's data needs be in five years and in 10 years? Is the proposed system sufficiently flexible to support future analysis and reporting needs?

Improve Accuracy

States can take several basic steps to improve the accuracy of their data. These aren't necessarily expensive, but they will provide a good return on investment.

Ensure that district personnel are trained and familiar with the system. Training of district data staff is a requirement for collecting reliable data. It is critical that the district data managers are thoroughly trained on the system and that they are cognizant of any changes that have been made to queries, procedures, schedules, or anything else that might affect data collection. The state should hold regular trainings and updates/refreshers for data managers in advance of data collections to ensure that everyone is prepared.

The state should develop a data dictionary and make it available to all users of the system. It should be kept up to date and users should be made aware any time it is revised. User guides and software manuals should document all features and procedures thoroughly and clearly. Finally, knowledgeable technical support should be made available on a regular schedule to support users of the system.

Automate the system as much as possible. Simplify the process of data entry as much as possible through the use of data-entry forms that include predefined options. Build in automated validity checks to the data-entry system that reduce key-entry errors.

Perform validity checks. Whatever the database or rate calculations that will be used, the process of importing data into a state enrollment database should involve a set of edits (validity checks) that will let data administrators ensure that they are receiving data of good quality. Some of these edits can be automated and others may be performed manually. They may be as simple as tallying the number of records submitted by each district and comparing that value to the same

value from the district's previous submission, or may involve more complex operations and comparisons.

Such edits will afford state data administrators a chance to discover and correct errors in the data before erroneous records are merged into the state database. The time required to perform these quality control checks is generally much less than would be required to discover and remedy errors caused by merging flawed data into an otherwise sound set of data. Westat's publication entitled "Taking Your Data to the Laundry: How to Clean Your Data" provides useful advice and information about collecting and preparing data for analysis.


Keep the data up to date. Districts must keep their databases as current as possible, making updates as the need arises. These data should be uploaded to the state's database at regular, frequent intervals.

Get your money's worth from your data. Data about the various indicators—graduation, dropout, parental involvement, transition, and post-school outcomes, for example—can help states identify districts' needs. At the local level, it can help schools and districts identify both systemic and student needs. Additionally, data should be leveraged to inform policy as well as the development and selection of school retention programs or interventions.

Since you have to collect the data, you might as well get your money's worth. Put it to work for more than just program evaluation! The data amassed in the development of SPPs is a good place to start. Look across the indicators for patterns that might indicate a need. Use this cross indicator strategy to help identify suitable programs or technical assistance strategies.

Summary

Steps that states and districts can take toward ensuring all parties involved in data collection, analysis, and reporting are using the same set of procedures, definitions, and calculations will increase the likelihood of calculating accurate dropout and graduation rates. This, in turn, will facilitate the identification of districts and schools that need assistance. It will also help SEAs and LEAs target



appropriate programs and services that address the factors contributing to dropout in these areas. Over time, this technical assistance will result in decreased dropout and an increase in the rate of school completion for youth with IEPs.

References

Brauen, M., Garrison-Mogren, R., Srivastava, S., Parshall, L. (2004). *Taking your data to the laundry: How to clean your data*. Rockville, MD: Westat.

Calculating graduation and dropout rates: A technical assistance guide. (1999). Prepared for the U.S. Department of Education, Office of Special Education Programs by Westat.

Data Accuracy: Critical Elements for Review of SPPs. Retrieved July 12, 2006 from <http://www.rrfcnetwork.org/content/view/283/47/>.

Graduation counts: A compact on state high school graduation data. (2005). Washington, DC: National Governors Association.

Schiller, E., O'Reilly, F., & Fiore, T. (2006). *The study of state and local implementation of the Individuals with Disabilities Act: Marking the progress of IDEA implementation*. Report to OSEP. Bethesda, MD: Abt Associates.

The National Dropout Prevention Center for Students with Disabilities' *SPP Tool Kit Series* are practical guides intended to assist state education agencies in their development and implementation of improvement activities associated with school completion. This guide is the first in the series.



The National Dropout Prevention Center for Students with Disabilities is funded by the U.S. Department of Education's Office of Special Education Programs Cooperative Agreement No. H326Q030002. The content therein does not necessarily reflect the views or policies of the U.S. Department of Education, nor does mention of other organizations imply endorsement by those organizations or the U.S. government.

National Dropout Prevention Center for
Students with Disabilities
Clemson University
209 Martin Street
Clemson, SC 29631-1555
Telephone: (800) 443-6392
TDD/TDY: (866) 212-2775
Fax: (864) 656-0136
Email: NDPCSD-L@clemson.edu
Web site: www.ndpc-sd.org